

Speech recognition for students with disabilities

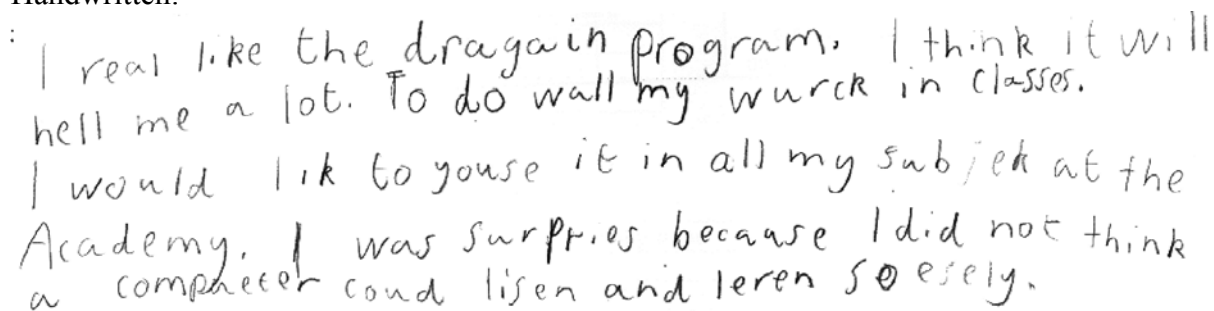
Introduction

“For so long they [dyslexic pupils] have had to rely on others to scribe for them that they have begun to despise themselves and ‘switch off’. VoiceText [DragonDictate] allows these pupils to respond independently, gives them control, and gives them the freedom to express themselves as others do.” (Donald, 1998)

“When I am doing the writing for him he co-operates, but when he writes himself (using voice recognition) he comes alive....” (Donegan, 2000a)

An evaluation from a student (Nisbet & Wilson, 2002):

Handwritten:



I real like the dragain program. I think it will
hell me a lot. To do wall my wurck in classes.
I would lik to youse it in all my subjek at the
Academy. I was surpris because I did not think
a compheer could lisen and leren so esely.

Dictated and then sent by email.....

“I really like the Dragon program. I think it will help me a lot. To do my work in the class/es. I would like to use it in all my subjects at the Academy. I was surprised because I did not think a computer could listen and learn so easily. I was very surprised when the words I was saying came up on the screen. I thought it was really excellent.” (Scott, age 12, reading age 8 years)

Speech recognition has been presented as the latest “panacea” for people with dyslexia and other writing difficulties. Some students have used speech recognition systems successfully for their work and for exams, and the use of this technology has helped them to overcome their difficulties and go on to higher education. However, there have been other students in other schools, who have been unsuccessful in using speech recognition. In this paper, we will try to describe the important factors to consider when introducing and using speech recognition with students with disabilities.

Speech recognition programs

Currently there are three main programs available for PC: DragonDictate, Dragon NaturallySpeaking and IBM ViaVoice. For the Macintosh, there are IBM ViaVoice and iListen. Each of the programs is sold in different variants, from basic, cheap versions for dictating text only, to more sophisticated programs for controlling the computer and dictating into different applications such as email. Refer to the web sites listed at the end of this paper for up to date information on currently available programs. In education generally, DragonDictate and NaturallySpeaking are more popular than ViaVoice.

Most of the programs are designed to recognise ‘continuous’ speech, where the user speaks in unbroken phrases or sentences. The exception is DragonDictate, where the user must leave a short gap (around 0.25 seconds) between each word. Although DragonDictate is an old program and its ‘word-by-word’, ‘discrete’ speech recognition may seem a less ‘natural’ technique, it has some advantages for some groups of students, as we shall discuss below.

All of the modern continuous speech recognition programs require a fairly powerful computer. For example, the minimum specification for running ViaVoice Pro is a Pentium 300 with 64 MB of RAM (for Windows 98) or 192 MB of RAM (Windows XP). Dragon NaturallySpeaking Preferred 7 requires a Pentium III 500 and 128 MB of RAM. However, these are the absolute minimum specifications and in practice the programs work much better with a faster processor and with more RAM. If you are buying a computer today the processor will be fast enough but you should buy 512 MB of RAM, particularly if you are running Windows XP or wish to dictate into applications like Microsoft Office.

Before using any speech recognition program, the user must first train it to recognise his or her voice, by reading a set text into the computer. A few years ago, this was time consuming and difficult but with the latest version of NaturallySpeaking, for example, the user has only to read 50 to 60 sentences which will take a good reader 5 to 10 minutes. The process of training a speech recognition program is no longer a barrier for most students and even students with severe visual or reading difficulties can be supported to get through the initial training. Students with speech impairments, due to cerebral palsy or dysarthria, for example, may have difficulties with the enrolment process on the modern ‘continuous’ programs if they cannot speak clearly or with ‘standard’ pronunciation. For these students, DragonDictate may be a better option because the training, although longer, appears to be more ‘forgiving’ for non-standard speakers.

Once the initial training is complete the program is ready for use and students with clear, articulate speech should be able to achieve more than 90% accuracy. However, since the program will always make mistakes the user must be able to review the text that has been dictated, identify errors and correct them. It is very important to correct recognition errors because this helps the speech recognition program improve its voice model for the user.

The table below attempts to summarise which programs are best suited to students with different skills.

The student has...	Consider ...
Average or above academic ability; moderate specific difficulties with reading and spelling	Dragon NaturallySpeaking Preferred or IBM ViaVoice
Average or above academic ability; moderate to severe specific difficulties with reading and spelling	Dragon NaturallySpeaking Preferred or DragonDictate with Keystone ScreenSpeaker. Keystone is able to read back the text as it is dated and also read out the words in the correction box.
Below average academic ability	DragonDictate with Keystone
Articulation difficulties.	DragonDictate
Average and above academic ability; difficulties with mouse or requires hands-free control	Dragon NaturallySpeaking Preferred or DragonDictate

Who might benefit from speech recognition?

Speech recognition can be helpful for people with physical access difficulties (e.g. repetitive strain injury, arthritis, high spinal injury) that make writing difficult (Donegan, 2000a). It can also be effective for students with reading, writing or spelling difficulties (e.g. dyslexia) and for those with visual impairment (Donald, 1998; Nisbet and Wilson, 2002; Litten, 2000; Higgins and Raskind, 2000; Miles, Martin & Owen, 1998). The references list more research reports and case studies about students with different types of difficulty that have used speech recognition.

There are no 'tests' that can predict whether a student will succeed with speech recognition – the only way to find out is to try it - but the main factors concerning the student are:

Age – Speech recognition programs at the present time do not usually work very well with children younger than about 9 or 10.

Cognitive skills – Writers must have reasonable general cognitive abilities in order to understand how the program works.

IT skills - Speech recognition programs are relatively complicated and so the student must be competent with the computer, or must be capable of learning how to operate it.

Motivation – Speech recognition systems can be frustrating to train and learn to use, so the writer must be very motivated *and supported* during the difficult initial stages.

Dictation skills – The essential skill needed for speech recognition is to be able to compose and dictate clear, well-structured sentences. Spoken language is different from written language, and students must learn how to speak 'written' English (or French, German, and hopefully Danish). While the programs are surprisingly good at coping with different accents, they cannot read minds or interpret poorly structured or grammatically incorrect language. The writer must also be able to speak reasonably distinctly and with good clarity.

Reading skills – The programs are never 100 % accurate and so the writer (or the writer and a helper) must be able to review the text that has been dictated, identify mis-recognised words, and then correct them. Most of the modern speech recognition programs have facilities to help with this process. Firstly, the student can highlight a piece of text and play back a recording of what was dictated. By comparing the recording with the text on screen, recognition errors can be identified. If the student has difficulty reading the text on screen then it can be highlighted and spoken out by the computer using text to speech feedback. The text-to-speech facilities provided by the speech recognition programs themselves are relatively basic and so if the writer has very poor reading skills it is worth considering a specialist text-to-speech program such as Keystone.

In fact, I believe that almost any student (apart from those with severe speech impairment or learning difficulties) can learn how to use speech recognition with some degree of success. The most likely reasons for the programs failing to work are more to do with technical difficulties and particularly, insufficient support from teaching staff. Speech recognition programs are complicated and learning to use them takes considerable time and effort -- particularly if the student has some form of physical, visual or learning difficulty.

Another major factor when discussing whether speech recognition is effective is the definition of "success". Some teachers, parents and students may be hoping that the speech recognition program will enable a struggling writer to produce large quantities of text as soon

as the student has finished the initial training. This is unlikely to happen because it takes considerable time and practice to learn how to compose and then dictate good written language. Students who have already learned how to use a dictaphone effectively are often more successful with speech recognition because they have already developed good dictating skills. Also, even if the programs were 100 % accurate and required no training whatsoever, many students would still prefer to type or use some other form of access technique or writing support. The question is whether speech recognition provides a faster, easier or more accurate method of writing than the alternatives and some students, having tried speech recognition and achieved good results, will decide that other methods of writing are more suited to them. In the CALL Centre project 72% of the students who went through the speech recognition training said that they would continue to use speech recognition for classwork. In most cases, the reasons for not using speech recognition were because other methods - such as word processing or using word prediction - were viewed to be more appropriate. Litten (Litten, 2000) reported that he expects "50 % or less of those I train to continue to use speech recognition, despite having selected them as individuals likely to benefit". Speech recognition is just another writing tool: it suits some writers for some writing tasks some of the time. Given that it does require considerable time and effort to learn to use the programs, teachers, parents and students should always consider whether other more basic writing methods (such as learning to touch-type) would be a more effective use of teaching and learning time and resources.

There is also a wider question about the educational effects of writing with speech recognition. Many researchers and teachers have reported that the use of speech recognition improves basic literacy skills, and the main reason for this is because speech recognition, when used carefully, provides a rich multisensory experience. The writer speaks the words, watches them appear on screen, and can review and listen to the text using speech output. In a ten-week trial of IBM Simply Speaking with eleven children (Miles, Martin & Owen 1998), the reading age of the pupils increased by an average of 13.4 months (British Ability Scales Reading Test) and the average spelling age by 6.1 months (Schonell test). However, this phenomena is unlikely to occur when writers dictate quickly without attending to the words as they appear. Many teachers and users of speech recognition therefore recommend the DragonDictate program because it forces the writer to dictate and check one word at a time. Elliott (Elliott, 1998) claims that "*continuous speech does not provide direct contact with individual words and is unlikely to facilitate the language improvements identified with discrete systems.*" She points out that "*discrete speech encourages writing at a slow rhythmical pace, which is calming, enabling clear thoughts. Spelling and reading also improve because every spoken word is spelled correctly and seen and heard simultaneously.*" While these statements about continuous and discrete speech recognition are not entirely supported by research (Miles et al (1998) found gains using continuous speech recognition, while Higgins and Raskind (2000) found both discrete and continuous programs produced improvements) there is no doubt that students with reading and writing difficulties must be taught to dictate slowly and carefully in order to develop their basic literacy using speech recognition.

There are also questions about writing style. Personally, I find it easier to compose good English when handwriting or typing than when using speech recognition. This is probably because of Elliott's point above - that handwriting and typing provide a direct contact with the words as they appear, whereas dictating complete sentences at the time does not provide the link and makes it more difficult to change the text as it is being written. For me, I can compose *while* handwriting or typing whereas when I use speech recognition I find that

composition must occur *prior* to dictating. With practice, I hope that this will change and that I will be able to dictate more fluently, but this issue is significant when introducing speech recognition to students.

Learning to use speech recognition

Most reports about speech recognition in education state that a key factor is adequate training and support. In the CALL Centre project, there was no clear correlation between the type of computer or the student's skills and whether they found speech recognition effective, whereas training and support was a major factor: "91% of the students who intend to continue using speech recognition received training once or more per week." (Nisbet, 2002). 63% of those who did not intend to continue using speech recognition had used it less than once a week.

Before introducing speech recognition to a student, the tutor must:

- consider which speech recognition program and other tools are most suitable for the student;
- learn how to use the program and become confident with it;
- go through the initial training *on the computer that the student will use* in order to check that the computer and speech recognition system are working correctly;
- set aside sufficient time to work with the student, ideally on an individual basis. The amount of time that this will take depends on the support that the student will require: able students with good IT and reading skills will require less support whereas less able students will require more support. The CALL Centre training pack comprises 10 individual lessons lasting between 30 minutes and 1 hour and most staff found they would work with the students individually for the first three or four lessons at least; Litten (2000) estimates that the most students require four training sessions totalling two hours; Donegan (2000b) provided between two and four hours support for students new to speech recognition.

When training a student to use speech recognition:

- the tutor should provide sufficient support to ensure that the session is successful (if necessary, by identifying and correcting any recognition errors for the student);
- the student must be drilled to always correct errors using the correction mechanism;
- the activities and exercises which student is asked to complete must be meaningful and productive;
- changes and improvements to the students speech files must be saved at the end of the session;
- the outcome of the session should be a hard copy that the student is happy with.

You can download complete copies of the CALL Centre speech recognition training materials for teachers and students from the CALL Centre web site:

<http://callcentre.education.ed.ac.uk/>

Summary

Speech recognition programs have been used successfully by students with a range of disabilities. The programs have been shown to provide an effective means of writing and recording work, and in some cases, have produced significant improvements in basic reading,

spelling and writing skills. Successful use of speech recognition requires considerable time, energy and commitment from teaching staff, management, parents and, especially, the student. Access to a suitable computer is also important. If this support cannot be provided, then alternatives, such as the use of a computer, word predictor, typing, or tape recorder may be more practical.

References

- BECTa *SEN Speech Recognition Project – Final Report*. BECTa. Available, with reports from schools at: <http://www.becta.org.uk/teachers/teachers.cfm?section=2&id=2142>
- Donald, E. (1998) *Voicetext: Evaluation of System when used by Students with Special Educational Needs*. Perth High School, Perth, Scotland. In *Introducing Speech Recognition in Schools CD*, CALL Centre, University of Edinburgh.
- Donegan, M. (2000a) *Voice Recognition Technology in Education: Factors for Success* Ace Centre Oxford. ISBN 1 903303 00 1.
- Donegan, M. (2000b) *BECTA Voice Recognition Project Report*. BECTA, <http://www.becta.org.uk/teachers/teachers.cfm?section=2&id=2142>
- Elliott E. (1998) *Breaking down the barriers*. Ability, 26, Autumn 1998. British Computer Society Disability Group.
- Higgins, L. E., Raskind, M.H., (2000) *Speaking to Read: the Effects of Continuous Versus Discrete Speech Recognition Systems on the Reading and Spelling of Children with Learning Disabilities*. Journal of Special Education Technology 15, 1. <http://jset.unlv.edu/15.1/higgins/first.html>
- Litten, M. *Usable Voice Recognition Software: The Mark College Experience!* Dyslexia Review, Vol. 13 No. 2. Also <http://www.keyspell.com/Speechproducts/speechproducts.htm>
- Nisbet, P.D. & Wilson, A. (2002) *Introducing Speech Recognition in Schools: using Dragon NaturallySpeaking*. Pub. CALL Centre, University of Edinburgh. ISBN 1 898042 22 X.
- Nisbet, P.D. & Wilson, A. (2002) *Introducing Speech Recognition in Schools: using IBM ViaVoice*. Pub. CALL Centre, University of Edinburgh. ISBN 1 898042 23 8.
- Nisbet, P.D. (2002) *Introducing Speech Recognition in Schools: Final Report to Scottish Executive Education Department*. CALL Centre, University of Edinburgh.
- Miles, Martin, Martin, Di & Owen, Jim (1998) *A Pilot Study into the Effects of Using Voice Dictation Software with Secondary Dyslexic Students*. Devon LEA.

Suppliers

- IANSYST, specialist UK suppliers, <http://www.dyslexic.com/>
- Keystone ScreenSpeaker, from Words Worldwide, <http://www.keyspell.com/>
- iListen Speech recognition for Macintosh. <http://www.macspeech.com/default.html>
- IBM ViaVoice. <http://www.scansoft.com/viavoice/>
- Dragon NaturallySpeaking. <http://www.scansoft.com/naturallyspeaking>

Paul D. Nisbet, CALL Centre, University of Edinburgh
August 2003